Scenario:

You are part of a team which is responsible for deploying applications to our public cloud production environment.

In the past the deployment process of the company consisted of a lot of manual steps which continuously led to human failures. Being fed up with the situation the CTO of the company entrusted you with the mission to automate this process and reduce the room for failure.

You received the following task:

* Enable the company to deploy their applications easily to a fully load balanced public cloud (e.g. AWS) from scratch.
* Use infrastructure as code to deploy (e.g. Ansible, Terraform)

On top of that he provided you with the following installation requirements for your environment:

* A docker container running the provided webservice

o Only use official images from docker hub

o Based on alpine-linux

o Oracle Java 8.x (not OpenJDK)

o To run the webservice execute: java -jar helloworld.war

(Service will be available on port 8080)

* A webserver (e.g. apache, nginx) which proxies http/s to the application container.

While creating all this, please keep all your scripts and configs under version control (e.g. Git) to keep a decent history of your work.

Our CTO is very keen to understand and follow how you got your job done.

Expected deliverables:

* All your dockerfiles, scripts, configs, playbooks, formulas etc. in a VCS (e.g. GitHub, Bitbucket).
* A short documentation to restore your files to be able to do the deployment based on your recipes.
* Do not provide pre-built docker images.

**Answer:**

As requested the exercises have been completed and all the config files are located on the github <https://github.com/juanpcamacho/smava>. As well I upload the docker image for the helloworld to my dockerhub repository. <https://hub.docker.com/r/jpcamach/smava/>

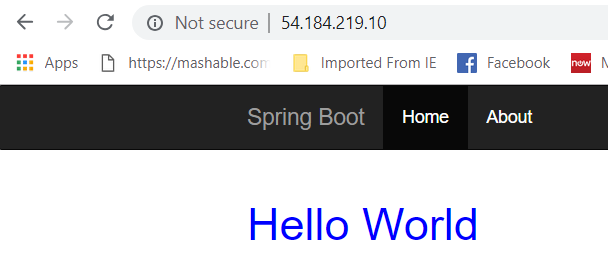
In order to enable the company to deploy their applications easily to a fully load balanced public cloud (e.g. AWS) from scratch we are going to use Kubernetes cluster managed by a Rancher orchestration with the hosts located in 2 different Availability Zones and Elastic Load Balancers to route and balance the traffic between the workers of the cluster. In this case we are going to assume we have our own Kubernetes available on AWS that can be shared between multiple applications at the same time. We are going to use Infrastructure as a Code for AWS and Ansible configuration manager. The ansible will oversee the application deployment with a playbook named ansible\_app\_deployment.yml that contain the following steps:

1. Creation of a ELB for load distribution between the workers of the cluster
2. Modification of a template that contain all the settings for the application deployment
   1. Persistence Volume and Persistence Volume Claim for Stateful data.
   2. Deployment including health checks for the pods.
      1. The pod have 2 containers the Hellowworld and NGINX
   3. Service for the DNS and exposure of the container.
3. Deployment of the Kubernetes via kubctl of the template mentioned above.
4. Copy of the file with the nginx configuration for the port forwarding.

To run the Ansible playbook you need to run the following command:

ansible-playbook -i hosts ansible\_app\_deployment.yml -e aws\_access\_key=<AWS\_ACCES\_KEY> -e aws\_secret\_key=<AWS\_SECRECT\_KEY> -e ELB\_NAME=smava-ELB -e app\_name=smava

At the end you should be able to see the application.



In case of any question please let me know!

Best Regards,